

FISHER MATRICES AND CONFIDENCE ELLIPSES: A QUICK-START GUIDE AND SOFTWARE

DAN COE

Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, MS 169-327, Pasadena, CA 91109
Version 1

ABSTRACT

Fisher matrices are used frequently in the analysis of combining cosmological constraints from various data sets. They encode the Gaussian uncertainties of multiple variables. They are simple to use, and I show how to get up and running with them quickly. Python software is also provided. I cover how to obtain confidence ellipses, add data sets, apply priors, marginalize, transform variables, and even calculate your own Fisher matrices. This treatment is not new, but I aim to provide a clear and concise reference guide. I also provide references and links to more sophisticated treatments and software.

Subject headings: cosmology

1. OUTLINE

I explain how to do/obtain the following with/from Fisher matrices:

- § 2: Confidence Ellipses
- § 3: Manipulation:
 - Marginalization, Priors, Adding Data Sets
 - § 4: How to Calculate your Own Fisher Matrices
 - § 5: How to transform variables
 - § 6: Dark energy pivot redshift
 - § 7: Discussion (brief) about what Fisher matrices are
 - § 8: Software I've come across (including my own)
 - § 9: How you can contribute to this paper

2. FISHER MATRICES \Rightarrow CONFIDENCE ELLIPSES

The inverse of the Fisher matrix is the covariance matrix:

$$[F]^{-1} = [C] = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \quad (1)$$

σ_x and σ_y are the $1-\sigma$ uncertainties in your parameters x and y , respectively (marginalizing over the other). $\sigma_{xy} = \rho\sigma_x\sigma_y$, where ρ is known as the correlation coefficient. ρ varies from 0 (independent) to 1 (completely correlated). Examples are plotted in Fig. 1.

The ellipse parameters are calculated as follows (e.g., Unknown 2008):

$$a^2 = \frac{\sigma_x^2 + \sigma_y^2}{2} + \sqrt{\frac{(\sigma_x^2 - \sigma_y^2)^2}{4} + \sigma_{xy}^2} \quad (2)$$

$$b^2 = \frac{\sigma_x^2 + \sigma_y^2}{2} - \sqrt{\frac{(\sigma_x^2 - \sigma_y^2)^2}{4} + \sigma_{xy}^2} \quad (3)$$

$$\tan 2\theta = \frac{2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2} \quad (4)$$

We then multiply the axis lengths a and b by a coefficient α depending on the confidence level we are interested in. For 68.3% CL (1- σ), $\Delta\chi^2 \approx 2.3$, $\alpha = \sqrt{\Delta\chi^2} \approx 1.52$. Other values can be found in Table 1. These can be calculated following e.g., Lampton et al. (1976).

Electronic address: coe@caltech.edu

TABLE 1
CONFIDENCE ELLIPSES:

σ	CL	$\Delta\chi^2$	α
1	68.3%	2.3	1.52
2	95.4%	6.17	2.48
3	99.7%	11.8	3.44

The area of the ellipse is given by

$$A = \pi(\alpha a)(\alpha b) \quad (5)$$

$$= \pi(\Delta\chi^2)ab \quad (6)$$

$$= \pi\sigma_x\sigma_y\sqrt{1 - \rho^2} \quad (7)$$

The inverse of the area is a good measure of figure of merit. The Dark Energy Task Force (DETF; Albrecht et al. 2006, 2009) used $FOM = \pi/A$ for the ability of experiments (WL, SN, BAO, CL) to constrain the dark energy equation of state parameters (w_0, w_a).

2.1. Probability $P(x, y)$

Interested in the probability that specific values are correct for parameters x and y ? The probability function $P(x, y)$ given best fit values (x_0, y_0) and $1-\sigma$ uncertainties (σ_x, σ_y) is calculated as follows:

$$\chi^2 = \frac{\left(\frac{\Delta x}{\sigma_x}\right)^2 + \left(\frac{\Delta y}{\sigma_y}\right)^2 - 2\rho\left(\frac{\Delta x}{\sigma_x}\right)\left(\frac{\Delta y}{\sigma_y}\right)}{1 - \rho^2} \quad (8)$$

$$P(x, y) = \exp\left(-\frac{\chi^2}{2}\right) \quad (9)$$

with $\Delta x \equiv x - x_0$ and $\Delta y \equiv y - y_0$. Note for $\rho = 0$ (uncorrelated x and y), the χ^2 formula looks familiar. For correlated x and y ($\rho > 0$), χ^2 is reduced.

3. MANIPULATION: MARGINALIZATION, PRIORS, ADDING DATA SETS, AND MORE

Consider a Fisher matrix provided by the DETF (Table 2) for optimistic Stage IV BAO observations for the following variables: $(\omega_m, \Omega_\Lambda, \Omega_k)$, where $\omega_m \equiv \Omega_m h^2$ and

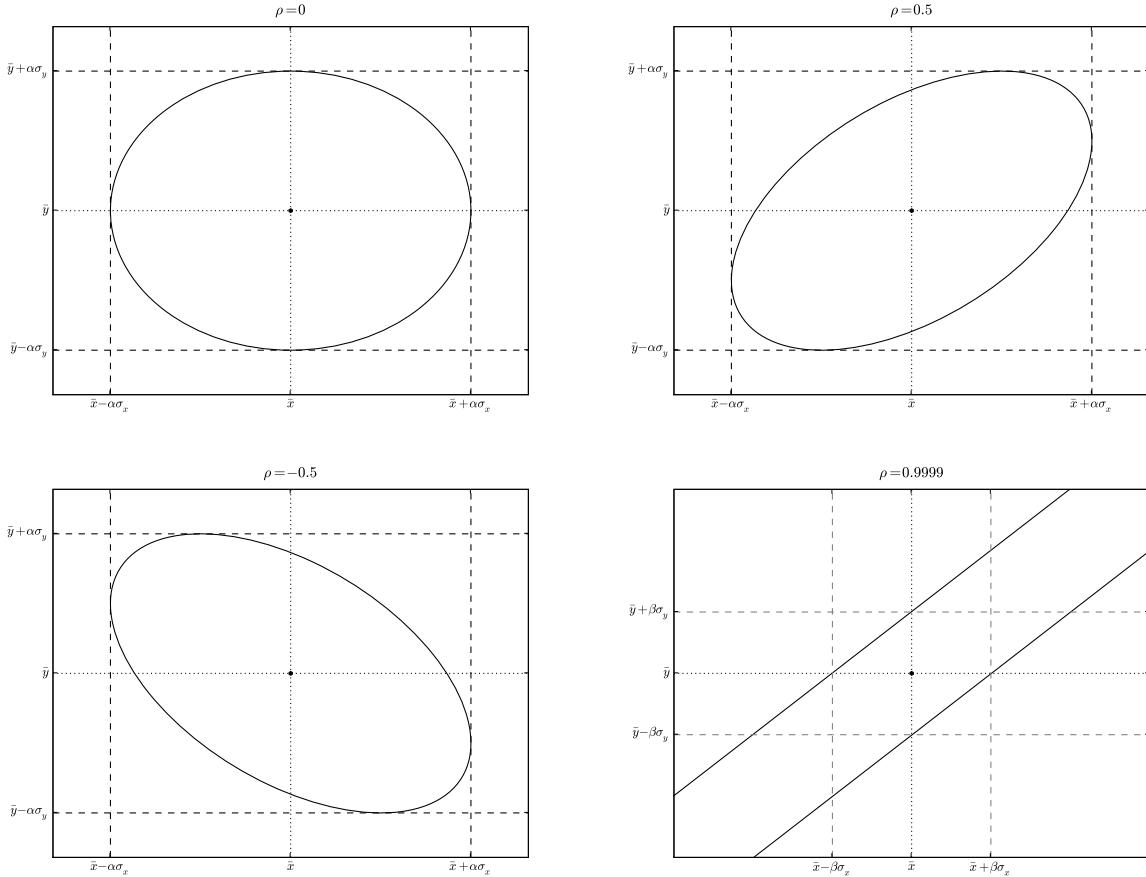


FIG. 1.— 68.3% (1- σ) confidence ellipses for parameters x and y with 1- σ uncertainties σ_x and σ_y and correlation coefficient ρ . In the first three panels, we plot as dashed lines the marginalized 1- σ uncertainty for each variable: $\alpha\sigma_x$ and $\alpha\sigma_y$, where $\alpha \approx \sqrt{2.3} \approx 1.52$. In the bottom-right panel, we zoom in to show the intersections with the axes: $\pm\beta\sigma_x$ and $\pm\beta\sigma_y$, where $\beta \approx 2.13\sqrt{1-\rho}$ (for $\rho \approx 1$).

TABLE 2
EXAMPLE FISHER MATRIX

	ω_m	Ω_Λ	Ω_k
ω_m	2,376,145	796,031	615,114
Ω_Λ	796,031	274,627	217,371
Ω_k	615,114	217,371	178,014

TABLE 3
CORRESPONDING COVARIANCE MATRIX

	ω_m	Ω_Λ	Ω_k
ω_m	3.20E-5	-1.56E-4	8.02E-5
Ω_Λ	-1.56E-4	8.71E-4	-5.25E-4
Ω_k	8.02E-5	-5.25E-4	3.69E-4

$\Omega_m + \Omega_\Lambda + \Omega_k = 1$. The covariance matrix (inverse of the Fisher matrix) is given in Table 3. For example, the top-left element tells us that $\Delta\omega_m \approx 0.00566 \approx \sqrt{3.20E-5}$.

3.1. Marginalization

When quoting these uncertainties on ω_m , the other variables (Ω_Λ, Ω_k) have automatically been marginalized over. That is, their probabilities have been integrated over: they have been set free to hold any values while we calculate the range of acceptable ω_m .

To calculate a new Fisher matrix marginalized over any

TABLE 4
FISHER MATRIX WITH FIXED $\Omega_k = 0$

	ω_m	Ω_Λ
ω_m	2,376,145	796,031
Ω_Λ	796,031	274,627

variable, simply remove that variable's row and column from the covariance matrix, and take the inverse of that to yield the new Fisher matrix.

3.2. Fixing Parameters

Suppose instead want the opposite: perfect knowledge of a parameter. For example, we want to consider a flat universe with a fixed value of $\Omega_k = 0$. To do this, simply remove Ω_k from the Fisher matrix (Table 4). The new covariance matrix and parameter uncertainties are calculated from the revised Fisher matrix.

Alternatively, the on-diagonal element corresponding to that parameter can be set to a very large value. For example, if we set the bottom-right element in Table 2 to 10^{12} , that would correspond to a 10^{-6} uncertainty in ω_m , or nearly fixed. Note that higher values in the Fisher matrix correspond to higher certainty.

3.3. Priors

Rather than fixing a parameter to an exact value, we may want to place a prior such as $\Delta\Omega_k = 0.01$ (1- σ).

In this case, simply add $1/\sigma^2 = 10^4$ to the on-diagonal element corresponding to that variable (in this case, the bottom left element).

3.4. Adding Data Sets

To combine constraints from multiple experiments, simply add their Fisher matrices: $F = F_1 + F_2$. Strictly speaking, any marginalization should be performed after the addition. But if the “nuisance parameters” are uncorrelated between the two data sets, then marginalization may be performed before the addition.

4. HOW TO CALCULATE YOUR OWN FISHER MATRICES

Given the badness of fit $\chi^2(x, y)$, your 2-D Fisher matrix can be calculated as follows:

$$[F] = \frac{1}{2} \begin{bmatrix} \frac{\partial^2}{\partial x^2} & \frac{\partial^2}{\partial x \partial y} \\ \frac{\partial^2}{\partial x \partial y} & \frac{\partial^2}{\partial y^2} \end{bmatrix} \chi^2 \quad (10)$$

In other words, $F_{ij} = \frac{1}{2} \frac{\partial \chi^2}{\partial p_i \partial p_j}$.

These derivatives are simple to calculate numerically:

$$\frac{\partial^2 \chi^2}{\partial x^2} \approx \frac{\chi^2(x_0 + \Delta x, y_0) - 2\chi^2(x_0, y_0) + \chi^2(x_0 - \Delta x, y_0)}{(\Delta x)^2} \quad (11)$$

$$\frac{\partial \chi^2}{\partial x} \approx \frac{\chi^2(x_0 + \Delta x, y_0) - \chi^2(x_0 - \Delta x, y_0)}{2\Delta x} \quad (12)$$

$$\frac{\partial^2 \chi^2}{\partial x \partial y} = \frac{\partial \frac{\partial \chi^2}{\partial x}}{\partial y} \quad (13)$$

5. TRANSFORMATION OF VARIABLES

Suppose we are given a Fisher matrix in terms of variables $p = (x, y, z)$ but we are interested in constraints on related variables $p' = (a, b, c)$. We can obtain a new Fisher matrix as follows:

$$F'_{mn} = \sum_{ij} \frac{\partial p_i}{\partial p'_m} \frac{\partial p_j}{\partial p'_n} F_{ij} \quad (14)$$

Let's spell this out explicitly. Here is the expression for element (a, b) in the new Fisher matrix:

$$F'_{ab} = \frac{\partial x}{\partial a} \frac{\partial x}{\partial b} F_{xx} + \frac{\partial x}{\partial a} \frac{\partial y}{\partial b} F_{xy} + \frac{\partial x}{\partial a} \frac{\partial z}{\partial b} F_{xz} \quad (15)$$

$$+ \frac{\partial y}{\partial a} \frac{\partial x}{\partial b} F_{yx} + \frac{\partial y}{\partial a} \frac{\partial y}{\partial b} F_{yy} + \frac{\partial y}{\partial a} \frac{\partial z}{\partial b} F_{yz} \quad (16)$$

$$+ \frac{\partial z}{\partial a} \frac{\partial x}{\partial b} F_{zx} + \frac{\partial z}{\partial a} \frac{\partial y}{\partial b} F_{zy} + \frac{\partial z}{\partial a} \frac{\partial z}{\partial b} F_{zz} \quad (17)$$

This can be calculated using matrices:

$$[F'] = [M]^T [F] [M] \quad (18)$$

where $M_{ij} = \frac{\partial p_i}{\partial p'_j}$:

$$[M] = \begin{bmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} & \frac{\partial x}{\partial c} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} & \frac{\partial y}{\partial c} \\ \frac{\partial z}{\partial a} & \frac{\partial z}{\partial b} & \frac{\partial z}{\partial c} \end{bmatrix} \quad (19)$$

and $[M]^T$ is the transpose.

All of these partial derivatives should be evaluated numerically, plugging in best-fit values of the parameters.

5.1. Transformation Example

Suppose we are given a Fisher matrix in terms of $(\omega_m, \Omega_\Lambda, \Omega_k)$, but we are interested in $(\Omega_m, \Omega_\Lambda, h)$. Here $\omega_m \equiv \Omega_m h^2$ and $\Omega_k = 1 - \Omega_m - \Omega_\Lambda$. Suppose further that the best-fit cosmology is $(\Omega_m, \Omega_\Lambda, h) = (0.3, 0.7, 0.7)$. Our transformation matrix is evaluated as follows:

$$[M] = \begin{bmatrix} \frac{\partial \omega_m}{\partial \Omega_m} & \frac{\partial \omega_m}{\partial \Omega_\Lambda} & \frac{\partial \omega_m}{\partial h} \\ \frac{\partial \Omega_\Lambda}{\partial \Omega_m} & \frac{\partial \Omega_\Lambda}{\partial \Omega_\Lambda} & \frac{\partial \Omega_\Lambda}{\partial h} \\ \frac{\partial \Omega_k}{\partial \Omega_m} & \frac{\partial \Omega_k}{\partial \Omega_\Lambda} & \frac{\partial \Omega_k}{\partial h} \end{bmatrix} \quad (20)$$

$$= \begin{bmatrix} h^2 & 0 & 2\Omega_m h \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 0.49 & 0 & 0.42 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix} \quad (22)$$

6. PIVOT REDSHIFT

Given the dark energy equation of state parameterization

$$w = w_0 + (1 - a)w_a \quad (23)$$

where $1/a = 1 + z$, if you have calculated a Fisher Matrix for dark energy parameters w_0 and w_a , go ahead and calculate the pivot redshift, too:

$$z_p = \frac{-1}{1 + \frac{\Delta w_a}{\rho \Delta w_0}} \quad (24)$$

At this redshift, $w(z)$ is best constrained (e.g., Fig. 16 of Huterer & Turner 2001). Rather than presenting constraints on (w_0, w_a) , constraints on (w_p, w_a) can be presented. That is, we constrain the value of w at $z = z_p$ rather than at $z = 0$ (along with w 's rate of change with time w_a).

The (w_p, w_a) confidence ellipse has no tilt; there is no correlation between the two, by definition.¹

¹ Thus the DETF chooses a more interesting ellipse to plot: (w_p, Ω_{DE}) .

But the area of the (w_p, w_a) ellipse is equal to the area of the (w_0, w_a) ellipse. From this and Eq. 7 it follows that

$$\Delta w_p = \Delta w_0 \sqrt{1 - \rho^2} \quad (25)$$

And if w is constant, then $\Delta w_p = \Delta w_0$.

Derivation of the pivot redshift formula follows from (Albrecht et al. 2006), calculating the uncertainty of

$$w_p = w_0 + (1 - a_p)w_a \quad (26)$$

$$(\Delta w_p)^2 = (\Delta w_0)^2 + ((1 - a_p)\Delta w_a)^2 + 2(1 - a_p)\Delta w_{0,a} \quad (27)$$

where $\Delta w_{0,a} = \rho\Delta w_0\Delta w_a$, and then minimizing Δw_p for a_p .

7. DISCUSSION

Fisher matrices encode the Gaussian uncertainties in a number of parameters. Confidence ellipses can be easily calculated over any pair of parameters. These provide an optimistic approximation to the true probability distribution. The true uncertainties may be larger and non-Gaussian. Note the best fit values themselves are not encoded in the Fisher matrices, and must be provided separately.

Fisher matrices allow one to easily manipulate parameter constraints over many variables. It is easy to add data sets, add priors, marginalize over parameters, and transform variables, as shown here.

A more in-depth discussion of Fisher matrices and issues surrounding their use can be found in (Albrecht et al. 2009).

This is the paper I'd wished I could find when I began my work with Fisher matrices: projections for cosmological constraints from gravitational lens time delays (Coe & Moustakas 2009).

² <http://www.its.caltech.edu/%7Ecoe/Fisher/>

³ <http://www.physics.ucdavis.edu/DETFast/>

⁴ <http://www.cosmology.org.za/>

8. SOFTWARE

Fisher.py² Python – simple manipulation of Fisher matrices and plotting of ellipses

DETFast³ (Albrecht et al. 2006) JAVA – Compare expectations of cosmological constraints from different experiments with your choice of priors with a few clicks!

Fisher4Cast⁴ (Bassett et al. 2009) Matlab – most sophisticated

Your ad here.

9. CONTRIBUTE

This is meant to be a brief guide, but if I've failed to reference another useful guide or your software or if I've neglected some detail (subtle or otherwise) about Fisher matrices, please e-mail me at coe(at)caltech.edu, and I'll be happy to update this document. Also please tell me if any section is unclear.

If I have not covered a useful topic, it is probably outside my knowledge of Fisher matrices. For example, I have not covered the analysis of Monte Carlo Markov Chains (MCMC) as provided, for example, by the WMAP Lambda website.⁵ If a generous reader could explain to me (or point me to an appropriate reference on) how to extract confidence contours and a Fisher matrix from a MCMC, I would be grateful and include the explanation here, giving due credit to the contributor.

I thank Olivier Dore for referring me to the DETFast software written by Jason Dick and Lloyd Knox whom I also thank for answering my questions about their software. It is a valuable resource. Once I took off these training wheels and began to produce my own plots, DETFast is still a valuable resource for Fisher matrices calculated by the DETF encoding their estimates of cosmological constraints from various future experiments.

This work was carried out at Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

⁵ <http://lambda.gsfc.nasa.gov/>

REFERENCES

Albrecht, A., Amendola, L., Bernstein, G., Clowe, D., Eisenstein, D., Guzzo, L., Hirata, C., Huterer, D., et al. 2009, ArXiv e-prints [ADS]

Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., et al. 2006, ArXiv Astrophysics e-prints [ADS]

Bassett, B. A., Fantaye, Y., Hlozek, R., & Kotze, J. 2009, ArXiv e-prints [ADS]

Coe, D. A. & Moustakas, L. A. 2009, ArXiv e-prints [ADS]

Huterer, D. & Turner, M. S. 2001, Phys. Rev. D, 64, 123527 [ADS]

Lampton, M., Margon, B., & Bowyer, S. 1976, ApJ, 208, 177 [ADS]

Unknown. 2008, Bivariate Normal Distribution and Error Ellipses [LINK]